THE

AMERICAN NATURALIST

VOL. XLVII *January, 1913* No. 553

FACTORS AND UNIT CHARACTERS IN MENDELIAN HEREDITY

PROFESSOR T. H. MORGAN

COLUMBIA UNIVERSITY

THE factorial hypothesis has played an important rôle in Mendelian heredity, and while students of Mendel's principles have had on the whole a pretty clear idea of the sense or senses in which they have made use of factors or symbols, yet those not engaged in the immediate work itself have, I believe, often been misled in regard to the meaning attached to the term factor, and by the assumed relation between a factor and a unit character. The confusion is due to a tendency, sometimes unintentional, to speak of a unit character as the product of a particular unit factor acting alone, but this identification has no real basis. It has, in fact, more than once been repudiated, yet the confusion has been so persistent that I venture to try to make clear my own position at least—it is one I think with which in the main many students of heredity will agree—in regard to the relation between unit-factors and unit-characters. I shall do this by means of several examples taken from my breeding experiments with the fly, *Drosophila ampelophila.*

The eye of this fly is red. A mutant arose with a vermilion eye. Crossed to the wild or red-eyed fly, the new color proved to be a Mendelian recessive.

According to the scheme that Mendel followed, red, *R*, and vermilion, *V*, are symbolized as complete and contrasting characters carried by the germ-plasm of the hybrid. They are assumed to separate in the germ-cells, and as a consequence two kinds of these cells are produced.

According to a more modern interpretation, known as the presence and absence theory, vermilion is supposed to arise through the loss of something from the germ-plasm of the wild fly. This something is not supposed to be the factor for vermilion, but another factor. On this scheme the red eye would be represented by the letters *RV*, and the vermilion eye by *rV*; as though the vermilion color arose through the loss of a red factor.

The relative advantages of these two modes of representation become apparent when two pairs of factors are involved. For instance, a new eye color—pink—appeared as a mutant. It, also, was recessive to red. Mendel's scheme would make the pink character the mate of the red character, just as vermilion had been before. But if pink and vermilion were mated to each other, it is not clear whether vermilion and pink should be treated as contrasted characters, or whether each should still be treated as allelomorphic to red. If either of these alternatives is adopted, the scheme fails to account for what actually happens. Mendel did not meet with such a situation, for none of his paired characters involved two changes in kind in the same organ, and consequently the problem did not exist for him.

Bateson did meet with just this situation in the case of the comb of fowls and the coat color of mice. His scheme, if applied to the present case of the eye colors in *Drosophila*, would be to represent red by *RV*, vermilion by *rV*, and pink by *Rv*. This scheme illustrates first why when vermilion is bred to pink a red-eyed fly, *rVRv*, should result; second, why in the second generation the proportion 9:3:3:1 should appear;[1] and third why in the eye

[1] Except in so far as modified by sex-linkage.

color series a new color is expected in the F_2 generation, represented here by *rv*. This new color I called orange, and since *rv* only meant two absences, I followed the conventional method and added the symbol *O* to stand for orange. The completed formulæ were:

RVO	red
rVO	vermilion
RvO	pink
rvO	orange

This is identical with the scheme that Bateson adopted for the mouse color series, viz:

GBCh	gray
gBCh	black
GbCh	cinnamon
gbCh	chocolate

In a later paper (1912) I used the symbol *P* instead of *R*, so that the series stood:

PVO	red
pVO	vermilion
PvO	pink
pvO	orange

Let us now examine some of the possible interpretations of these symbols to see in what sense the letters were used for factors.

It is undoubtedly *implied, on the presence and absence scheme,* that something is *lost* from the original germplasm *PVO* when the vermilion *pVO* arises. The vermilion color is supposed to be the product of what is left when this something (called *P*) is lost. It is not supposed on this hypothesis that the vermilion factor alone is responsible for the vermilion color, for it is hypothetically only a part of what is left when something (*P*) is lost. Yet it is the identification of the vermilion factor with the vermilion eye-color that the opponents of Mendelism seem anxious to impute to the Mendelians.

Again, when the pink eye mutant appeared, it would have been assumed, on the presence and absence theory, that something was lost, so that the formula is *PvO*. Here again the pink color is the result of all that is left when something (*V*) is lost. Pink is not assumed to be produced by a factor *P*, but by what is left when a factor *V* is lost. An egg is supposed to have lost something and vermilion developed, another egg is assumed to have lost something else and pink developed. It was the loss of the vermilion factor that allowed pink color to develop, and the loss of the pink factor that allowed vermilion color to develop. When pink and vermilion are mated together, the original color—red—is restored, because on this scheme what each has lost is made good by what is found in the other.

To my series of eye color factors the letter *O* was added to indicate the nature of the color produced when two factors, *P* and *V*, were assumed to be absent. The symbol *O* at that time did not seem to stand in the formulæ on the same footing as *P* and *V*, because it stood for a color, and not for a factor that had been lost from the germ-cells of the wild fly. But since on the presence and absence scheme *O* stood for the residuum after *P* and *V* were lost it stood for the same sort of thing as did *P* and *V*, for *P* and *V* also stood for residua, *when they were not used as symbols for factors*. This will be made clearer later.

When the experiments had progressed to this stage, a new eye color appeared that was called eosin. Mated to orange it gave red; therefore, it seemed that this mutant must have contained *P* and *V*, and I inferred that it owed its color to the loss of an imaginary *O* factor. Eosin was represented, therefore, by *PVo*. But a moment's thought will show that on this scheme, *as long as P and V are present, any* loss from the germ-plasm (giving a new eye color) added to orange should give red, because orange would contain what the new mutant had lost.

The history of this case will show how, with the best of

intentions, one may be led into a paradoxical position in regard to the use of factors. Even admiting that the representation is purely symbolic, the letters used may unintentionally come to stand for different things. Thus in the case first cited, the letter P in the formula vP stood for a *residuum* that gave pink, but in the formula Vp, the letter p stood for the loss of a *P-factor,* yet p is the allelomorph of P, which latter, as stated, meant the residuum when V was lost. In other words, a double meaning was attached to P, for it stood both for the P-factor, which was only a part of the residuum, and also for the residuum as a whole. It is this doubleness of meaning that gives the opponents of Mendelian inheritance an occasion to impose upon the factorial hypothesis a meaning that is really foreign to it. Admitting that the Mendelians themselves have not always taken the pains to state explicitly that the symbols represent both a factor and a residuum, there is still little or no justification in imputing to the presence and absence theory the view that a given character, pink color, for instance, is the product of a pink factor alone. The attempt to impute to the factorial hypothesis the same interpretation that Weismann made use of in his theory of determinants rests largely upon an erroneous understanding of the symbolism employed. Weismann identifies each character of the organism as the product of a special determinant. The factorial hypothesis assumes only that the cell in one case is different from the cell in the other, the difference relating, it is true, to some part, but the character produced may be the result of the whole or of much of the cell, and not of one part alone.

There is a further difference between these two points of view. A change in a factor may have far-reaching consequences. Every part of the organization capable of reacting to the new change is affected. Though we seize upon the most conspicuous difference between the old type and its mutant, and make use of this alone, every student of heredity is familiar with cases where more

than the part taken as the index is affected. Weismann's theory, on the other hand, seems as a rule to identify each character with a special determinant for that character, and his meaning is clear when it is remembered that the process of development on Weismann's view is a process of sorting out of the determiners of the germ-plasm into different regions of the body. The factorial hypothesis makes no such assumption, but refers differentiation to the interaction of the parts on each other—every cell retaining the full complex of the original germ-plasm. Hence the possibility of the far-reaching effects of any change in the germ-plasm!

II

The presence and absence *system of nomenclature* (aside from its implications as to what is meant by presence and absence) has till the present time justified itself, when properly interpreted, by its usefulness. It seems to me that as a system of nomenclature it may be used, if one so desires, quite apart from the idea, that a loss in a character involves necessarily a loss in the germ-plasm. I can bring forward one clear case at least that seems to me difficult to explain if absence is taken literally to mean the loss of a factor from the germ complex. I refer to a mutation "backwards," which in the older terminology meant reversion, or atavism.[2] In my pure cultures (at rare intervals) individuals have appeared like the original progenitors of the stock. I have not scrupled to put aside this evidence, because contamination, even with extreme care, will occasionally occur; and even if a reversion had occurred there would be no way of proving that it was such and not contamination. In fact, eosin first appeared in white-eyed stock and seemed to arise through reversion, but at the same time it seemed so improbable that this could happen that I tried to account for its appearance in a roundabout way. Now I should say that the factor *w* reverted to *W*.

[2] It is needless to add, perhaps, that atavism by recombination is not here for a moment brought into question.

But the clear case referred to above is the following: Quite recently there appeared in a culture bottle that had been producing for more than four months (probably for twelve generations) only wingless flies, an individual with one "wingless" wing and one normal wing on the other side. Here the evidence is conclusive that reversion had occurred. The wingless stock in which the asymmetrical form arose had purple eyes and the same eye color was present in the new type. As the eye color was relatively new at the time the chance that contamination had occurred was rendered very unlikely. Had contamination by a red-eyed fly occurred, making the new type a heterozygote, the eye color would have been the dominant red. When the asymmetrical fly (♂) was bred to wingless females only wingless flies appeared, for three or more generations. The reversion, therefore, was somatic and did not involve the germ-plasm, yet this fact does not invalidate the question here raised.

In the light of this evidence, as well as the evidence from ever-sporting varieties (that may also be considered, I think, as mutating and reverting as regular processes), I believe it unwise to commit ourselves any longer to a view that a recessive character is necessarily the result of a loss from the germ-cell. We need only assume that some readjustment occurs, and as the result a new factor is produced. A simile may make this clearer, if not taken too literally. If we suppose that a factor is a labile aggregate, and that a rearrangement in it occurs, then the new aggregate in connection with the other parts of the cell produces a character that differs from the old one. Here there need be no loss, but only a change in configuration with a corresponding change in the end product in which the changed part plays a rôle, along with the other parts of the cell. A factor, in this sense, may exist in two or more forms according to the state of equilibrium; one of its states is dominant-producing, and the other is recessive-producing. Such a view may make it easier for us to appreciate that a mutation need

not be a loss, and that a recessive may revert in the sense that it may mutate. In chemical terms, the process is reversible.

III

As I have pointed out, the presence and absence nomenclature, if properly understood, offers no practical difficulties so long as only two changes in the same organ are involved, but in experiments with *Drosophila* we have passed beyond this stage and must have at command a system by means of which more than two factors may be easily and conveniently represented. How impossible it becomes to use the presence and absence nomenclature when new characters are appearing may be shown by the following illustrations.

As already stated, Mendel's method of representing the allelomorphic pairs sufficed so long as one new character is contrasted with the original one. In this sense the relation of a vermilion-eyed mutant to the red-eyed fly could be fully represented by treating red (*R*) and vermilion (*V*) as allelomorphs. But when another mutation in eye color appeared the scheme was no longer feasible. Now, in the same sense in which it became necessary to supplant Mendel's scheme by another one, it becomes necessary to change the presence and absence scheme when a third mutation appears in the same organ; for, the presence and absence scheme is not sufficiently elastic to allow the introduction of a new term in the series, unless a complete revision of the method is made each time that a new mutation in kind occurs.

For example, when it becomes desirable to compare the eosin eye with the vermilion-pink (or orange eye already known) it becomes puzzling to know what symbols to adopt. If, as I assumed, the symbol *O* in *VPO* is changed to small *o*, then the formula for eosin becomes *VPo*. But this is inconsistent with the scheme already adopted because the small letter *o* stands for a character called eosin. If to avoid this ambiguity a letter *E* (or *e*) is introduced for eosin the situation is even more puzzling.

The only logical method that could be followed, if an attempt is made to apply consistently the current scheme of presence and absence[3] would be the following:

When it becomes necessary to construct a series, let us say one involving three characters (*PVE*), the three double recessives (*Pve, pVe, pvE*) must be made up and suitable names given to them, the initial letters of these names then become the factors sought. Such a procedure not only involves holding in suspense the naming of the factors until all the double recessives have been obtained, but involves renaming all the factors, each time a new series is made up.[4] This method is not likely to recommend itself if a simpler one can be employed. The plan here advocated avoids such difficulties.

The first letter (or the first and second or some other significant letter) of the name of the new character stands, as heretofore, as its symbol; thus *P* stands for the pink factor and small *p* stands for the correlative factor of the pink-eyed fly. Whether small *p* represents the loss of the *P* factor, or a change in that factor when the pink eye appears, is immaterial. The large letter represents the dominant character in conformity with the current scheme.[5] The eye color series will then be:

Red	*PVE*
Vermilion	*PvE*
Pink	*pVE*
Vermilion-pink	*pvE*
Eosin	*PVe*
Eosin-vermilion	*Pve*
Eosin-pink	*pVe*
Eosin-pink-vermilion	*pve*

[3] It is the nomenclature that is here brought into question and not, for the moment, the underlying conception of presence and absence, for even in my scheme this conception might still be held if it seemed desirable to do so.

[4] When, as in the case of the mouse colors, all the members of the series are known, there is no difficulty in finding suitable symbols, for the current names of the characters give the letters for the symbols.

[5] When a new *dominant* character appears it is represented by the capital letter and its allelomorph in the original form by a small letter.

The same scheme might be followed by using the small letters for the factors in the original red eye: thus red = *pve;* and the capital letter for the corresponding factor in the mutant; thus, vermilion = *pVe,* etc. A disadvantage of this scheme is that the large letter now stands for a recessive condition and the small letter (its allelomorph) for a dominant condition. Usage has, however, made us accustomed to interpreting a large letter as a dominant, its corresponding recessive (its allelomorph) by a small letter and therefore the plan first suggested seems more desirable. It is with much reluctance that I suggest this change in our present nomenclature. It has become necessary, however, in the case of the *Drosophila* to find some way to represent consistently those cases in which three or more factors are involved in the same organ. The change is not one of any theoretical importance, but a practical necessity for all cases of this kind. When one new character is contrasted with the original one, Mendel's way may still be the simplest and easiest way of formulating the results, and will, no doubt, be followed. When two new characters are involved the formula of presence and absence is a sufficient way of representing the symbols. But when new mutations are appearing some other plan must be adopted. The one here suggested has at least two merits: it is as easy to use as either of the foregoing for one and for two characters, and can also be utilized when any number of further mutations appear in the same organ.

The scheme applied to body colors is as follows: Two mutants arose, yellow and black, and by recombination, a "brown" or yellow-black fly was obtained. The symbols would be wild fly = *YYBB,* yellow = *yyBB,* black *YYbb,* and yellow-black *yybb.* Two other mutations in body color have appeared, both dark, one is called ebony, *eb* and the other sable, *s.* When brought in connection with the preceding mutation the gametic symbols would be:

Wild fly	YBE_bS
Yellow	yBE_bS

Black	YbE_bS
Yellow-black	ybE_bS
Ebony	YBe_bS
Sable	YBE_bS
Etc.	

Another combination is represented by certain wing mutations. A mutant called miniature appeared and may be represented by *m;* another mutant appeared, called rudimentary, and may be represented by *r;* and a third form, produced by recombination, was called miniature-rudimentary, *mr.* The symbols for this series would be:

Wild fly	*MR*
Miniature	*mR*
Rudimentary	*Mr*
Rud.-min.	*mr*

Later several other mutations in wings also appeared. Six of these may be selected for illustration, viz: Vestigial,[6] v_g; Bifid, b_i; Arc, a_r; Curved, c_v; Jaunty, *j;* Balloon, b_a. If these are brought into connection with the foregoing the symbols for the wild fly in terms of these factors would be:

$$\text{Wild Fly} = M,\ R,\ V_g,\ B_i,\ A_r,\ C,\ J,\ B_a.$$

In order to study the relation of these characters to each other it has become necessary to combine many of them and in order to represent the results some system of symbols must be adopted. Obviously, it would be highly undesirable to be obliged to revise the system each time that any of these new mutations are brought into relation with those that have already been compared.

It may be asked, why may not the current scheme be retained, since in most cases only two characters are likely to be involved and characters can always be contrasted in pairs on this scheme? The answer is that more

[6] This is the wingless fly of former papers.

characters in the same organ have already been obtained, and it is at least as important to have a scheme by which they can be represented as it is to have a scheme where two characters only are studied. Moreover, the more characters that are obtained that show association in inheritance the further we may hope to go in our analysis of the consitution of the germ-plasm, which is admittedly the fundamental problem in the study of heredity. We must have some convenient way of representing the symbols in order to carry out this analysis, and on the grounds of convenience alone some scheme other than the current one must be found, at least for such a case as this of *Drosophila*. Another scheme has, in fact, been adopted by Baur and Hagedoorn. The letters that stand for the factors bear no relation to the name of the characters involved. This scheme allows the addition of any number of new factors to a series under consideration. In practise, however, this plan makes it extremely difficult to understand what any formula means without continual reference to the key of symbols used. We have found in practise that the scheme is so puzzling when several factors are under consideration that we have been led to follow the current method of representing each factor by the initial letter (or other suggestive letters) of the character that it stands for. Except in this regard the method of formulation here suggested is similar in principle to the A.B.C. scheme of Baur.